

## Molecular basis for the folding of $\beta$ -helical autotransporter passenger domains

Yuan, Xiaojun; Johnson, Matthew D; Zhang, Jing; Lo, Alvin W; Schembri, Mark A; Wijeyewickrema, Lakshmi C; Pike, Robert N; Huysmans, Gerard H M; Henderson, Ian R; Leyton, Denisse L

DOI:

[10.1038/s41467-018-03593-2](https://doi.org/10.1038/s41467-018-03593-2)

License:

Creative Commons: Attribution (CC BY)

*Document Version*

Publisher's PDF, also known as Version of record

*Citation for published version (Harvard):*

Yuan, X, Johnson, MD, Zhang, J, Lo, AW, Schembri, MA, Wijeyewickrema, LC, Pike, RN, Huysmans, GHM, Henderson, IR & Leyton, DL 2018, 'Molecular basis for the folding of  $\beta$ -helical autotransporter passenger domains', *Nature Communications*, vol. 9, no. 1, 1395. <https://doi.org/10.1038/s41467-018-03593-2>

[Link to publication on Research at Birmingham portal](#)

### **Publisher Rights Statement:**

Yuan, X., Johnson, M. D., Zhang, J., Lo, A. W., Schembri, M. A., Wijeyewickrema, L. C., Pike, R. N., Huysmans, G. H. M., Henderson, I. R., & Leyton, D. L. (2018) Molecular basis for the folding of  $\beta$ -helical autotransporter passenger domains. *Nature Communications* 9. 1395. <https://doi.org/10.1038/s41467-018-03593-2>

### **General rights**

Unless a licence is specified above, all rights (including copyright and moral rights) in this document are retained by the authors and/or the copyright holders. The express permission of the copyright holder must be obtained for any use of this material other than for purposes permitted by law.

- Users may freely distribute the URL that is used to identify this publication.
- Users may download and/or print one copy of the publication from the University of Birmingham research portal for the purpose of private study or non-commercial research.
- User may use extracts from the document in line with the concept of 'fair dealing' under the Copyright, Designs and Patents Act 1988 (?)
- Users may not further distribute the material nor use it for the purposes of commercial gain.

Where a licence is displayed above, please note the terms and conditions of the licence govern your use of this document.

When citing, please reference the published version.

### **Take down policy**

While the University of Birmingham exercises care and attention in making items available there are rare occasions when an item has been uploaded in error or has been deemed to be commercially or otherwise sensitive.

If you believe that this is the case for this document, please contact [UBIRA@lists.bham.ac.uk](mailto:UBIRA@lists.bham.ac.uk) providing details and we will remove access to the work immediately and investigate.

## ARTICLE

DOI: 10.1038/s41467-018-03593-2

OPEN

# Molecular basis for the folding of $\beta$ -helical autotransporter passenger domains

Xiaojun Yuan<sup>1</sup>, Matthew D. Johnson<sup>1</sup>, Jing Zhang<sup>1</sup>, Alvin W. Lo<sup>2,3</sup>, Mark A. Schembri<sup>2,3</sup>,  
Lakshmi C. Wijeyewickrema<sup>4</sup>, Robert N. Pike<sup>4</sup>, Gerard H.M. Huysmans<sup>5</sup>, Ian R. Henderson<sup>6</sup> &  
Denisse L. Leyton<sup>1,7</sup>

Bacterial autotransporters comprise a C-terminal  $\beta$ -barrel domain, which must be correctly folded and inserted into the outer membrane to facilitate translocation of the N-terminal passenger domain to the cell exterior. Once at the surface, the passenger domains of most autotransporters are folded into an elongated  $\beta$ -helix. In a cellular context, key molecules catalyze the assembly of the autotransporter  $\beta$ -barrel domain. However, how the passenger domain folds into its functional form is poorly understood. Here we use mutational analysis on the autotransporter Pet to show that the  $\beta$ -hairpin structure of the fifth extracellular loop of the  $\beta$ -barrel domain has a crucial role for passenger domain folding into a  $\beta$ -helix. Bioinformatics and structural analyses, and mutagenesis of a homologous autotransporter, suggest that this function is conserved among autotransporter proteins with  $\beta$ -helical passenger domains. We propose that the autotransporter  $\beta$ -barrel domain is a folding vector that nucleates folding of the passenger domain.

<sup>1</sup>Research School of Biology, Australian National University, Canberra, ACT 0200, Australia. <sup>2</sup>School of Chemistry and Molecular Biosciences, The University of Queensland, Brisbane, QLD 4072, Australia. <sup>3</sup>Australian Infectious Disease Research Centre, The University of Queensland, Brisbane, QLD 4072, Australia. <sup>4</sup>Department of Biochemistry and Genetics, La Trobe Institute for Molecular Science, La Trobe University, Melbourne, VIC 3086, Australia. <sup>5</sup>Department of Physiology and Biophysics, Weill Cornell Medicine, New York, NY 10065, USA. <sup>6</sup>Institute of Microbiology and Infection, University of Birmingham, Birmingham B15 2TT, UK. <sup>7</sup>Medical School, Australian National University, Canberra, ACT 0200, Australia. These authors contributed equally: Xiaojun Yuan, Matthew D. Johnson. Correspondence and requests for materials should be addressed to D.L.L. (email: [denisse.leyton@anu.edu.au](mailto:denisse.leyton@anu.edu.au))

Type V autotransporter proteins (currently classified as types Va to Ve) are a distinct class of outer membrane proteins (OMPs) that share a common secretion pathway, which facilitates their navigation through the bacterial cell envelope<sup>1</sup>. Although the different subtypes of type V autotransporters are distinguished by variations in their domain organization, oligomerization state, and structure, they are unified by a similar transmembrane topology<sup>1</sup>. Embedded in the outer membrane is a  $\beta$ -barrel domain that forms a secretion pore, which is needed for the translocation of the (often covalently linked) passenger domain to the bacterial cell surface<sup>1</sup>.

The prototypical type Va (“classical”) autotransporter consists of a 12-stranded C-terminal  $\beta$ -barrel domain and an N-terminal passenger domain<sup>2</sup> that, in most cases, is folded or predicted to fold into an elongated  $\beta$ -helix<sup>3–6</sup>, although globular folds have also been reported<sup>7</sup> and predicted to occur<sup>6</sup>. Once folded, passenger domains have key roles in pathogenesis; some passenger domains remain attached to the  $\beta$ -barrel to function as adhesins, whereas others are processed and released into the environment to function as toxins, or to mediate disruption of the host immune response<sup>8–11</sup>.

In a cellular context, key molecules catalyze classical autotransporter folding and membrane insertion as they do for other types of  $\beta$ -barrel proteins. The translocation and assembly module participates in autotransporter assembly<sup>12–14</sup>, as do periplasmic chaperones such as Skp and SurA<sup>15–18</sup>. Based on evidence from cross-linking studies that show autotransporters directly interact with BamA, the central component of the  $\beta$ -barrel assembly machinery<sup>15,19,20</sup>, it has been proposed that the passenger domain is translocated through a pore formed by the transient fusion of the autotransporter and BamA  $\beta$ -barrels during membrane insertion. BamA structures and molecular dynamics simulations show conformational flexibility along the seam that hydrogen bonds the first with the last  $\beta$ -strand of the BamA  $\beta$ -barrel<sup>21,22</sup>. However, it is unclear whether BamA can transiently form fusion pores with substrate OMPs.

Irrespective of the composition of the translocation pore, there is a general agreement that in classical autotransporters, the passenger domain is translocated in a C- to N-terminal direction where the most C-terminal  $\beta$ -helical rungs fold into a stable, protease-resistant structure called the ‘stable core’<sup>3,19,23,24</sup>. It has been proposed that folding of the portion of the passenger domain that appears first at the cell surface into a stable scaffold nucleates processive folding of the more N-terminal segments of the passenger domain during translocation. In support of this hypothesis are chemical<sup>3,24,25</sup> and mechanical<sup>26</sup> denaturation studies showing that the core is the most stable part of the protein. Moreover, several lines of evidence suggest that stacking ( $\pi$ - $\pi$ ) interactions between buried aromatic residues in the stable core not only provide stability to this region, but also drive translocation and vectorial folding of the passenger domain<sup>19,26,27</sup>. Although collectively these results provide insight into how the passenger domain is translocated and folded into its functional form, many mechanistic details are lacking.

Here, using the classical Pet autotransporter, an enterotoxin from enteroaggregative *Escherichia coli*<sup>28</sup> as a model protein, we present evidence that the fifth  $\beta$ -hairpin loop (L5), which connects  $\beta$ -strands 9 and 10 at the extracellular surface of the  $\beta$ -barrel domain, templates folding of the passenger domain. By monitoring the assembly of Pet variants carrying modifications in the L5  $\beta$ -hairpin in live bacterial cells, we provide evidence that  $\beta$ -strand propensity of the loop, but not its specific amino acid sequence is required for folding of the passenger domain into a stable  $\beta$ -helix. Moreover, biophysical studies on refolded protein suggest that the nucleation of passenger domain folding can occur independently of the cellular machinery, and is entirely mediated

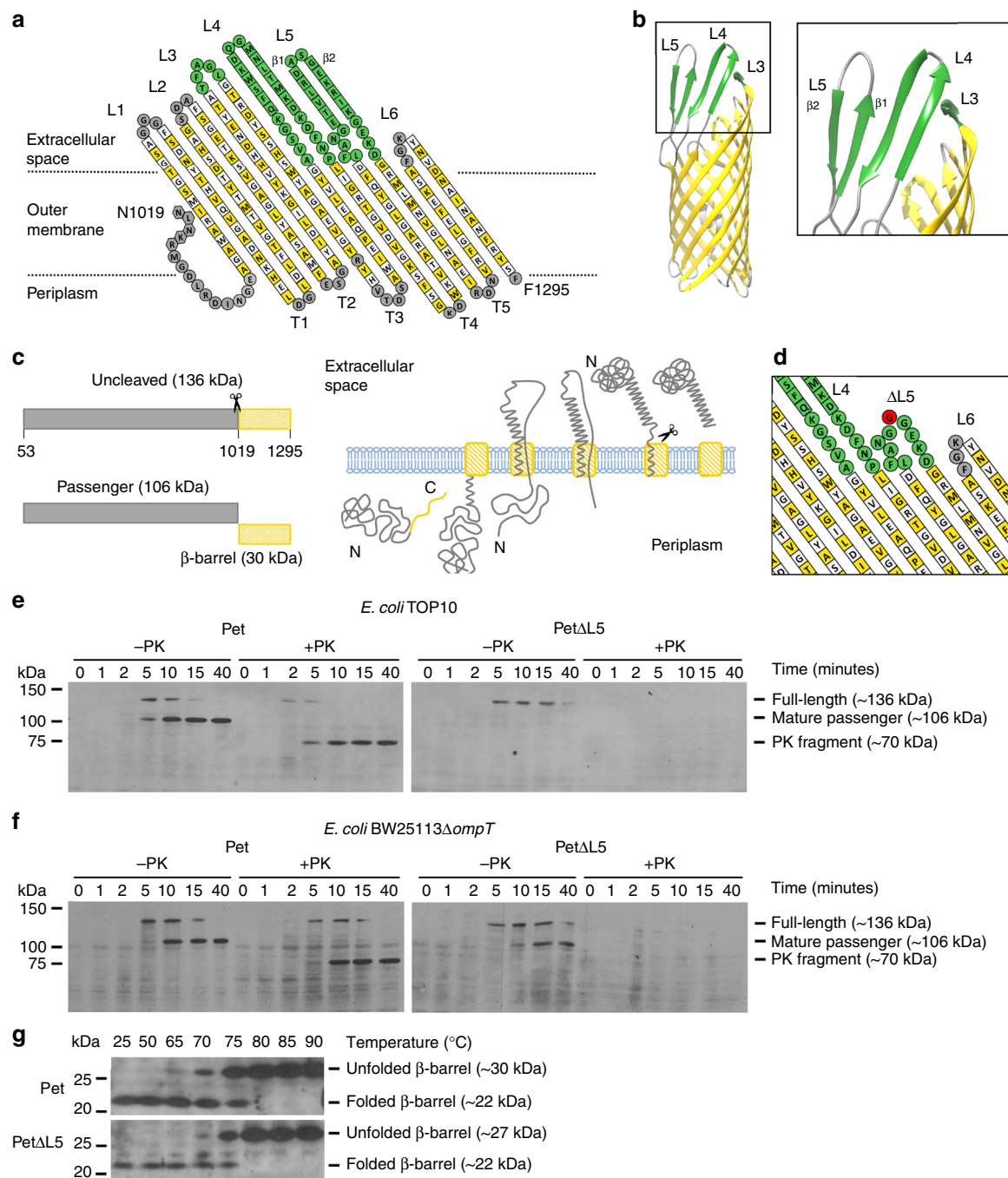
by the  $\beta$ -barrel domain. Experimental analysis of a second classical autotransporter, EspP, and in silico sequence and structure comparison with other autotransporters lead us to suggest that L5-assisted folding of the passenger domain represents a conserved folding mechanism for classical autotransporter proteins.

## Results

**Folded Pet is largely resistant to proteinase K digestion.** The homologous autotransporter EspP from enterohaemorrhagic *E. coli* has been crystallized<sup>29,30</sup>, thus providing a means to model the topology of the Pet  $\beta$ -barrel domain (Fig. 1a) and to infer its three-dimensional structure (Fig. 1b). Pet and EspP belong to the SPATE (serine protease autotransporters of *Enterobacteriaceae*) subfamily of autotransporters, which are characterized by an autocatalytic processing of the passenger from the  $\beta$ -barrel between two conserved asparagine residues in the intra-barrel  $\alpha$ -helical segment that connects both domains<sup>30–32</sup> (Fig. 1a, b). Processing of the full-length ~136 kDa Pet and EspP pro-proteins results in the secretion of the ~106 kDa mature form of the passenger domains into the culture supernatant, whereas the ~30 kDa  $\beta$ -barrel domains remain embedded in the outer membrane<sup>28,33</sup> (Fig. 1c). We have established a time-course assay for the pulse-chase expression of Pet in whole cells that allows its assembly to be monitored by SDS-polyacrylamide gel electrophoresis (PAGE) and immunoblotting using antibodies that selectively recognize the Pet passenger domain<sup>34</sup>. Moreover, the addition of proteinase K to a duplicate sample at each time-point provides information about the state of folding of the molecule<sup>34</sup>. Indeed, while periplasmic proteins are protected from proteinase K digestion, unless the outer membrane is first permeabilized by polymyxin B<sup>34</sup> (Supplementary Fig. 1a, b), the passenger is partially digested into a ~70 kDa protease-resistant fragment when surface exposed and natively folded<sup>34</sup> (Fig. 1e, left panel, + PK).

## Truncation of extracellular L5 perturbs Pet assembly in vivo.

Truncation of a long  $\beta$ -hairpin loop (L4) that connects  $\beta$ -strands 7 and 8 at the extracellular surface of the  $\beta$ -barrel domain of BrkA, a classical autotransporter from *Bordetella pertussis*, was shown to reduce the extent of passenger domain translocation<sup>35</sup>. Hence, we wanted to test whether the longer extracellular loops of the  $\beta$ -barrel domain (Fig. 1a, b, green) also have a role in the assembly of SPATEs. Thus, we truncated the loops that connect  $\beta$ -strands 9 and 10 (L5; Pet $\Delta$ L5) (Fig. 1d), 7 and 8 (L4; Pet $\Delta$ L4), and 5 and 6 (L3; Pet $\Delta$ L3) (Supplementary Fig. 2a), and compared the assembly of these mutant proteins with the assembly of wild type Pet using the pulse-chase expression assay<sup>34</sup>. Consistent with previous results<sup>34</sup>, processing of full-length Pet (which here refers to Pet minus its signal peptide) into distinct passenger and  $\beta$ -barrel fragments was in process by 5 min and largely completed by 10 min (Fig. 1e, left panel, – PK). Proteinase K digestion of Pet into a ~70 kDa protease-resistant fragment on the same timescale showed that the mature passenger domain molecules were natively folded (Fig. 1e, left panel, + PK). In contrast, the ~106 kDa mature Pet $\Delta$ L5 passenger was absent at all time points (Fig. 1e, right panel,  $\pm$  PK). Furthermore, although full-length Pet $\Delta$ L5 was detected by the 5 min time point (Fig. 1e, right panel, – PK), complete digestion of the passenger domain molecules by proteinase K was observed (Fig. 1e, right panel, + PK), suggesting that they were surface exposed in an unfolded or severely misfolded conformation. The presence of cleaved and similarly heat-modifiable Pet and Pet $\Delta$ L5  $\beta$ -barrels in total membrane preparations, detected using antibodies that selectively recognize the Pet  $\beta$ -barrel domain (Fig. 1g), suggests that truncation of L5 does not inhibit folding or insertion of the  $\beta$ -barrel into bacterial outer membranes, or autocatalytic release of the passenger



**Fig. 1** Truncation of L5 affects folding of the Pet passenger domain in bacterial cells. Topology model of the Pet  $\beta$ -barrel domain (**a**) based on the crystal structure of the EspP  $\beta$ -barrel domain (PDB code 3SLJ) (these  $\beta$ -barrels share ~91% sequence identity) (**b**). The structure shows L5 in an 'open' conformation where it interacts with L4, forming a 4-stranded  $\beta$ -sheet. It is noteworthy that the uncleaved passenger domain that traverses the length of the pore is hidden for clarity. For **a**, yellow and white squares indicate residues with side chains that point toward the lipid bilayer and the  $\beta$ -barrel lumen, respectively. Residues in periplasmic turns labeled T1-T5 and extracellular loops labeled L1, L2, and L6 are shown in gray, as are the residues in the  $\alpha$ -helical segment. Residues in extracellular loops labeled L3, L4, and L5 are shown in green. Given that these loops are mobile, the alignment of L4 and L5 does not strictly follow PDB 3SLJ. **c** The domain structure (left) and schematic for biogenesis (right) show that, post cytoplasmic synthesis and inner membrane translocation, Pet is processed by a signal peptidase to a ~136 kDa polypeptide in the periplasm. The  $\beta$ -barrel domain of Pet is then assembled into the outer membrane to allow translocation of the passenger domain, in a C- to N- terminal direction, until the extreme N-terminus traverses the  $\beta$ -barrel pore. An autocatalytic cleavage reaction within an intra-barrel  $\alpha$ -helical segment liberates the ~106 kDa mature form of the passenger domain, leaving the ~30 kDa  $\beta$ -barrel domain embedded in the outer membrane. **d** Topology model of the Pet  $\beta$ -barrel domain showing the L5 truncation created by replacing E<sub>1233</sub>-K<sub>1246</sub> with one glycine residue (shown in red). Pulse-chase expression of Pet and Pet $\Delta$ L5 and sensitivity to proteinase K (PK) in *E. coli* TOP10 (**e**) and BW25113  $\Delta$ ompT (**f**) monitored by SDS-PAGE and immunoblotting with anti-Pet passenger domain antibodies. All samples were TCA precipitated prior to SDS-PAGE. **g** Heating of total membranes containing Pet and Pet $\Delta$ L5 at the temperatures indicated. Samples were analyzed by SDS-PAGE and immunoblotting with anti- $\beta$ -barrel domain antibodies. Images are representative of at least two independent experiments



domain into the culture supernatant. These data are consistent with a previous study on EspP where Barnard et al.<sup>29</sup> aimed to investigate the contribution of L5 on  $\beta$ -barrel domain stability. In addition to finding that truncation of L5 in EspP ( $\Delta E_{1238-E_{1253}}$ ) had no effect on the stability of this domain, the authors also noted no effect on passenger domain translocation or cleavage.

To account for the disappearance of the ~106 kDa mature Pet $\Delta$ L5 passenger, pulse-chase expression of Pet and Pet $\Delta$ L5 was repeated in *E. coli* BW25113 $\Delta ompT$ <sup>36</sup>, a strain lacking the outer membrane protease OmpT. OmpT digests unfolded passenger domains on the bacterial surface, while natively folded passengers are resistant to OmpT degradation<sup>37</sup>. As expected, pulse-chase expression assays showed that the assembly of Pet in *E. coli* TOP10 (Fig. 1e, left panel,  $\pm$  PK), BW25113 $\Delta ompT$  (Fig. 1f, left panel,  $\pm$  PK), and wild-type BW25113 (Supplementary Fig. 3, left panel,  $\pm$  PK) occurred in a similar fashion. The assembly of Pet $\Delta$ L5 in TOP10 (Fig. 1e, right panel,  $-$  PK) and BW25113 in the absence of proteinase K (Supplementary Fig. 3, right panel,  $-$  PK) also occurred in a similar manner. In contrast, there were clear differences in the assembly of Pet $\Delta$ L5 in BW25113 $\Delta ompT$  in the absence of proteinase K (Fig. 1f, right panel,  $-$  PK). In BW25113 $\Delta ompT$ , full-length Pet $\Delta$ L5 was visible by the 5 min time point where processing into distinct passenger and  $\beta$ -barrel fragments occurred in a manner similar to wild type Pet (Fig. 1f, left panel,  $-$  PK). These data suggest that in TOP10 and BW25113, the Pet $\Delta$ L5 passenger is autocatalytically cleaved from its  $\beta$ -barrel domain, yet is unable to fold into its native conformation and is therefore susceptible to degradation by OmpT. Moreover, the presence of the full-length ~136 kDa species in all three strains suggests that when attached to its  $\beta$ -barrel, the passenger domain is less accessible to OmpT proteolysis (Fig. 1e, f and Supplementary Fig. 3, right panels,  $-$  PK). Nevertheless, complete degradation of the ~136 kDa and ~106 kDa species observed after proteinase K treatment (Fig. 1e, f and Supplementary Fig. 3, right panels,  $+$  PK) demonstrates that the Pet $\Delta$ L5 passenger, whether in its full-length form and still attached to the L5-truncated  $\beta$ -barrel or cleaved and released into the culture supernatant as a mature protein, is unable to fold into its native conformation when exposed on the bacterial surface. Barnard et al.<sup>29</sup> did not report a similar defect in passenger folding upon L5-truncation in EspP. However, the EspP variant used by these authors [EspP $\Delta$ 1( $\Delta$ L5)] only contained the 116 residues of the passenger domain immediately N-terminal to the L5-truncated  $\beta$ -barrel domain where the small passenger size (~13 kDa) would make it challenging to accurately pick up folding defects using pulse-chase assays<sup>29</sup>.

There was no major difference in the folding of Pet $\Delta$ L3 relative to that observed in wild type Pet, where the generation of a ~70 kDa protease-resistant fragment in samples treated with proteinase K suggested that the passenger domain was able to fold into its native conformation (Supplementary Fig. 2b,  $+$  PK). In contrast, the cleavage of Pet $\Delta$ L4 occurred with substantially reduced efficiency, with very little processing of the 136 kDa species to the 106 kDa mature species after 40 min (Supplementary Fig. 2b,  $-$  PK). The 136 kDa form of Pet $\Delta$ L4 was only degraded by proteinase K in the presence of polymyxin B, suggesting that most of the passenger is exposed to the periplasm (Supplementary Fig. 2c,  $+$  polymyxin B and PK). However, the low translocation efficiency of Pet $\Delta$ L4 makes it difficult to assess whether L4 also has a role in passenger folding (Supplementary Fig. 2b,  $+$  PK). Pulse-chase expression of Pet $\Delta$ L4 in BW25113 $\Delta ompT$  visualized greater amounts of the 136 kDa and 106 kDa species (Supplementary Fig. 2d, e,  $-$  PK), suggesting that some of the translocated Pet $\Delta$ L4 passenger domain is susceptible to OmpT degradation and is thus unfolded. The low translocation efficiency of Pet $\Delta$ L4 was rescued in a mutant with a partially

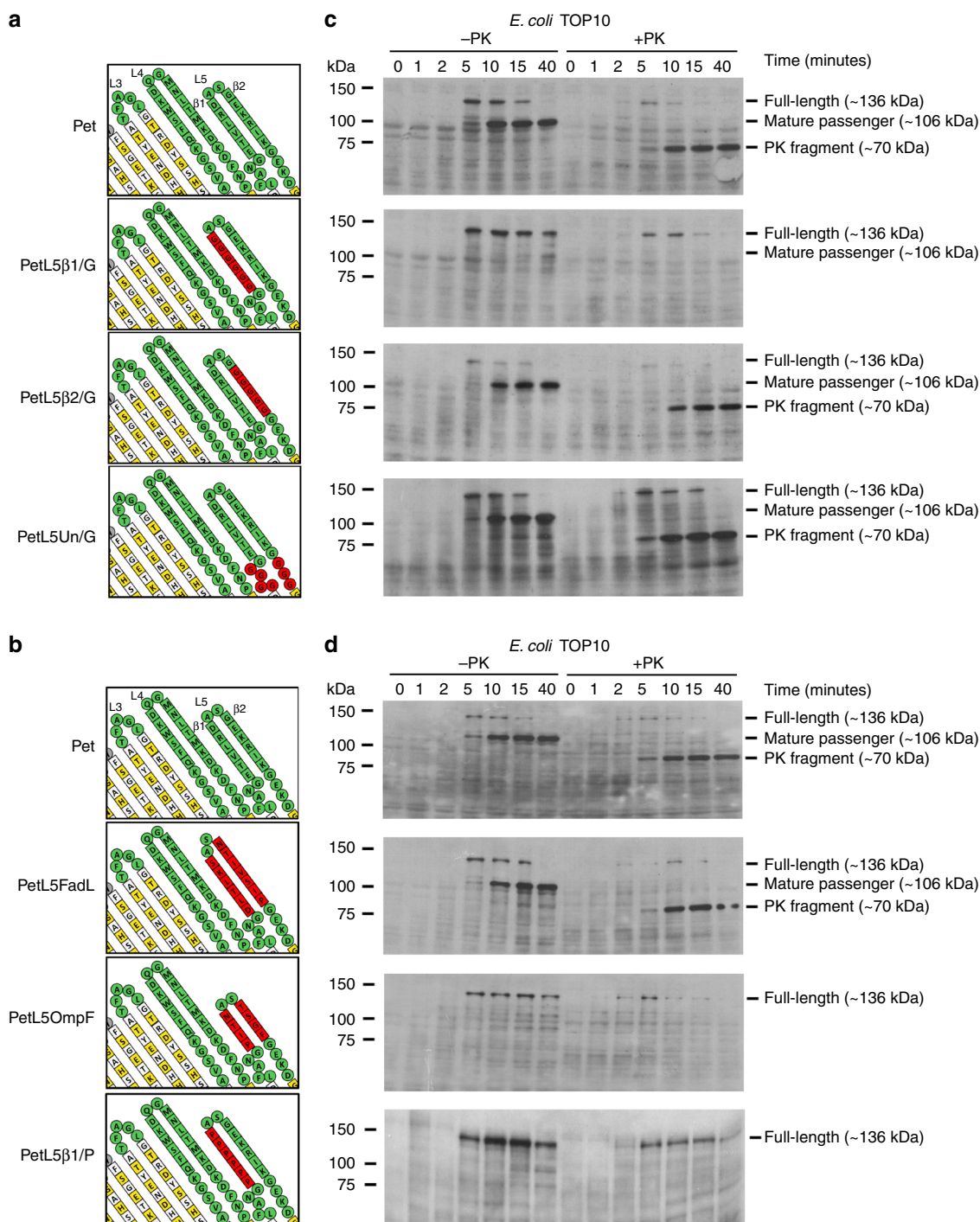
truncated L4  $\beta$ -hairpin, Pet $\Delta$ L4P (Supplementary Fig. 2f). Pet $\Delta$ L4P displayed a concomitant delay in the arrival of the passenger domain at the cell surface and autocatalytic processing into distinct passenger and  $\beta$ -barrel domain fragments relative to that observed in wild type Pet (Supplementary Fig. 2g,  $\pm$  PK). Generation of a ~70 kDa protease-resistant fragment in samples treated with proteinase K showed that the passenger domain of Pet $\Delta$ L4P was able to fold into its native conformation despite the delay in its surface exposure (Supplementary Fig. 2g,  $+$  PK). Collectively, these data indicate that the L4  $\beta$ -hairpin is mainly required for passenger translocation to occur efficiently, as shown previously for a L4 truncation variant of BrkA<sup>35</sup>. Furthermore, while L4 could have a role in passenger folding, either directly or indirectly, it appears not to be essential for this process.

### Passenger folding relies on the $\beta$ -hairpin conformation of L5.

To identify the amino acid residues of L5 responsible for mediating passenger folding, Pet variants were created (Fig. 2a) and analyzed by pulse-chase expression. Pet<sup>L5 $\beta$ 1/G</sup>, a variant with all amino acids in the first  $\beta$ -strand in L5 ( $E_{1233-D_{1238}}$ ) mutated to glycine, demonstrated a severe defect in passenger domain folding similar to that observed for Pet $\Delta$ L5 (Fig. 2c). As a small portion of the 136 kDa form of Pet<sup>L5 $\beta$ 1/G</sup> was not degraded by proteinase K (Fig. 2c,  $+$  PK), we performed pulse-chase expression of Pet<sup>L5 $\beta$ 1/G</sup> in BW25113 $\Delta ompT$  to test whether a translocation defect is the explanation for the absence of the ~70 kDa protease-resistant fragment that is a marker for correct passenger folding. In BW25113 $\Delta ompT$ , full-length Pet<sup>L5 $\beta$ 1/G</sup> was visible by the 5 min time point where processing into distinct passenger and  $\beta$ -barrel fragments occurred in a manner similar to wild-type Pet (Supplementary Fig. 4a, b,  $-$  PK), whereas treatment with proteinase K failed to yield a prominent ~70 kDa protease-resistant fragment (Supplementary Fig. 4b,  $+$  PK). These data confirm that in TOP10, the Pet<sup>L5 $\beta$ 1/G</sup> passenger is autocatalytically cleaved from its

$\beta$ -barrel domain, yet is unable to efficiently fold into its native conformation and is therefore susceptible to degradation by OmpT. In contrast, there were no obvious differences in the folding of Pet<sup>L5 $\beta$ 2/G</sup>, a variant with all amino acids in the second  $\beta$ -strand in L5 ( $E_{1242-K_{1246}}$ ) mutated to glycine, relative to that observed in wild type Pet (Fig. 2c). Pet<sup>L5 $\beta$ 3/G</sup>, a variant with all amino acids in the unstructured region beneath the  $\beta$ -hairpin ( $L_{1228-N_{1231}}$  and  $E_{1248-D_{1250}}$ ) mutated to glycine, was similarly unaffected (Fig. 2c). Together, these data suggest that a stretch of six amino acids ( $E_{1233-D_{1238}}$ ) in the first  $\beta$ -strand in L5 mediate folding of the Pet passenger domain.

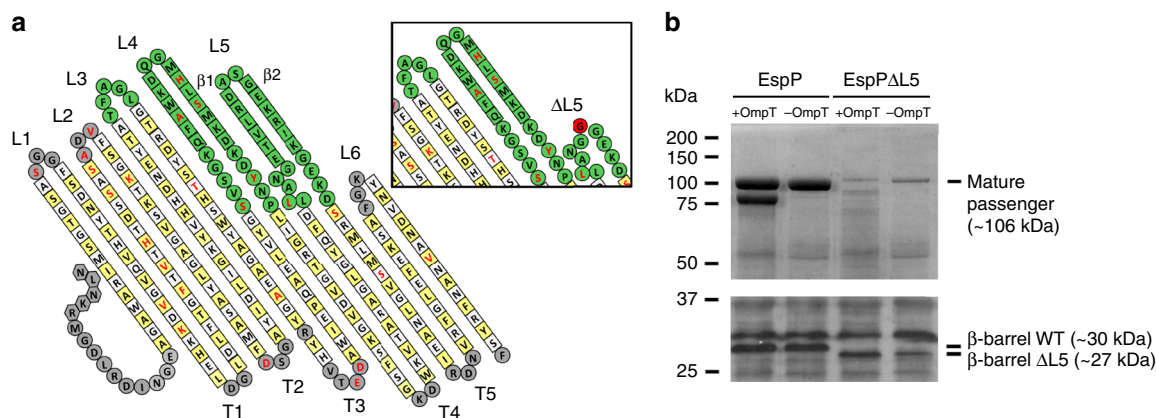
To determine whether the side-chains of the six residues in first  $\beta$ -strand in L5 facilitate passenger folding by providing specificity to the interaction, a series of site-directed mutants were created (Supplementary Fig. 5a) and analyzed by pulse-chase expression. There were no obvious differences in the folding of mutants in which all hydrophobic residues were mutated to the polar amino acid asparagine (Pet<sup>L5 $\beta$ LI/N</sup>) or in which charged residues were swapped to the opposite charge (Pet<sup>L5 $\beta$ SR/D,E/K</sup>) to disrupt potential salt bridges with the second  $\beta$ -strand (Supplementary Fig. 5c), relative to that observed in wild type Pet (Supplementary Fig. 5b). Finally, there was no obvious difference in the folding of Pet<sup>L5 $\beta$ TD/A</sup>, a mutant with the two remaining residues in the first  $\beta$ -strand in L5 ( $T_{1234}$  and  $D_{1238}$ ) mutated to alanine, relative to that observed in wild-type Pet (Supplementary Fig. 5a, b). The tolerance of amino acids with dissimilar physicochemical properties in all of these positions suggests that the nucleation of passenger folding does not depend on specific amino acid side-chain interactions. Rather, passenger folding might still occur, because the substitutions do not disrupt backbone-backbone



**Fig. 2** L5 likely templates folding of the Pet passenger domain by  $\beta$ -strand augmentation. Topology model of the Pet  $\beta$ -barrel domain showing the Pet<sup>L5 $\beta$ 1/G</sup>, Pet<sup>L5 $\beta$ 2/G</sup>, and Pet<sup>L5Un/G</sup> (**a**), and the Pet<sup>L5FadL</sup>, Pet<sup>L5OmpF</sup>, and Pet<sup>L5 $\beta$ 1/P</sup> (**b**) mutations (in red). Pulse-chase expression of Pet, Pet<sup>L5 $\beta$ 1/G</sup>, Pet<sup>L5 $\beta$ 2/G</sup>, and Pet<sup>L5Un/G</sup> (**c**), and of Pet, Pet<sup>L5FadL</sup>, Pet<sup>L5OmpF</sup>, and Pet<sup>L5 $\beta$ 1/P</sup> (**d**), and sensitivity to proteinase K (PK) in *E. coli* TOP10 monitored by SDS-PAGE and immunoblotting with anti-Pet passenger domain antibodies. All samples were TCA precipitated prior to SDS-PAGE. Images are representative of at least two independent experiments

hydrogen bond interactions between the passenger and L5, nor the  $\beta$ -sheet structure of the L5  $\beta$ -hairpin. This hypothesis is consistent with our findings demonstrating that substitution of the residues in the first  $\beta$ -strand in L5 with a stretch of six glycine residues severely perturbs folding of the passenger domain (Fig. 2a, c) conceivably through disruption of the  $\beta$ -sheet structure in this region (the average solution conformation of a six residue glycine-based peptide is a  $3_1$  extended helix<sup>38,39</sup>). To

further validate whether disruption of the  $\beta$ -sheet structure perturbs passenger folding, we substituted all amino acids in the first  $\beta$ -strand in L5 for proline to create Pet<sup>L5 $\beta$ 1/P</sup>, eliminating potential hydrogen bonding partners required for the formation of a  $\beta$ -sheet (Fig. 2b). Similar to Pet<sup>L5 $\beta$ 1/G</sup>, Pet<sup>L5 $\beta$ 1/P</sup> had a severe defect in passenger folding (Fig. 2d), strongly supporting the hypothesis that  $\beta$ -strand structure in L5 is required for efficient passenger folding. As in Pet<sup>L5 $\beta$ 1/G</sup>, a small portion of the 136 kDa



**Fig. 3** Truncation of L5 affects folding of the EspP passenger domain. **a** Left panel: Topology model of the EspP  $\beta$ -barrel domain based on its crystal structure (PDB code 3SLJ). Yellow and white squares indicate residues with side chains that point toward the lipid bilayer and the  $\beta$ -barrel lumen, respectively. Residues in periplasmic turns labeled T1-T5 and extracellular loops labeled L1, L2, and L6 are shown in gray, as are the residues in the  $\alpha$ -helical segment. Residues in extracellular loops labeled L3, L4, and L5 are shown in green and residues that differ in Pet are highlighted in red font. Right panel: topology model of the EspP  $\beta$ -barrel domain showing the L5 truncation created by replacing E<sub>1238</sub>-K<sub>1251</sub> with one glycine residue (shown in red). **b** SDS-PAGE of TCA-precipitated culture supernatant fractions (top panel) and immunoblotting of whole-cell lysates with anti-Pet  $\beta$ -barrel domain antibodies\* (bottom panel) collected after growth of *E. coli* TOP10 (+ OmpT) and BW25113 $\Delta$ ompT (- OmpT) expressing EspP and EspP $\Delta$ L5. Image is representative of at least two independent experiments. \*The Pet and EspP  $\beta$ -barrel domains are so similar that anti-Pet  $\beta$ -barrel domain antibodies cross-react with the EspP  $\beta$ -barrel domain

form of Pet<sup>L5 $\beta$ 1P</sup> was not degraded by proteinase K (Fig. 2d). Therefore, we performed pulse-chase expression of Pet<sup>L5 $\beta$ 1P</sup> in BW25113 $\Delta$ ompT, which again ruled out a translocation defect as an explanation for the absence of a prominent ~70 kDa proteinase-resistant fragment (Supplementary Fig. 4a, b,  $\pm$  PK).

Next, we replaced the residues E<sub>1233</sub>-D<sub>1238</sub> and G<sub>1241</sub>-K<sub>1246</sub> in L5 with residues D<sub>316</sub>-S<sub>321</sub> and N<sub>324</sub>-P<sub>330</sub> from loop 5 from the long-chain fatty acid transporter FadL, where similar to L5 in Pet the first  $\beta$ -strand in the  $\beta$ -hairpin is six residues in length, and with residues P<sub>239</sub>-N<sub>242</sub> and T<sub>247</sub>-F<sub>250</sub> from loop 6 from the porin OmpF, which has a shorter four-residue  $\beta$ -strand 1 to create Pet<sup>L5FadL</sup> and Pet<sup>L5OmpF</sup>, respectively (Fig. 2b and Supplementary Fig. 5d). There was no obvious difference in the folding of Pet<sup>L5FadL</sup> relative to that observed in wild type Pet (Fig. 2d). In contrast, pulse-chase expression of Pet<sup>L5OmpF</sup> demonstrated a severe defect in passenger domain folding similar to that observed for Pet $\Delta$ L5 (Fig. 2d). Yet again, pulse-chase expression of Pet<sup>L5OmpF</sup> in BW25113 $\Delta$ ompT (Supplementary Fig. 4a and b,  $\pm$  PK) confirmed that, in TOP10, the mature Pet<sup>L5OmpF</sup> passenger domain is largely unfolded and therefore susceptible to degradation by OmpT. One explanation for this discrepancy is that the heterologous OmpF loop does not adopt a  $\beta$ -hairpin when grafted onto the Pet  $\beta$ -barrel domain. However, if the heterologous OmpF loop does adopt a  $\beta$ -hairpin conformation when grafted onto the Pet  $\beta$ -barrel domain, then these data suggest that  $\beta$ -strand propensity and the length of the L5  $\beta$ -hairpin, but not its specific amino acid sequence, have a role in the efficient folding of the Pet passenger domain.

#### L5-assisted folding could be a conserved folding mechanism.

To test whether L5-assisted folding of the passenger domain is specific to Pet, we examined the influence of the L5 truncation on the SPATE EspP<sup>29,30</sup> by truncating the same amino acid residues in L5 of EspP as in Pet to create EspP $\Delta$ L5 (Fig. 3a). The Pet and EspP  $\beta$ -barrel domains are 91% identical (Supplementary Fig. 6b) and have L5 regions that differ by only one amino acid in 23 residues, where a F<sub>1229</sub> to L substitution is found in the unstructured region beneath the  $\beta$ -hairpin (Fig. 3a). In contrast,

although both passenger domains fold into an elongated  $\beta$ -helix<sup>40,41</sup>, their sequence identity is 50%. Hence, EspP is ideally suited to investigate whether L5-assisted folding of the passenger domain could represent a conserved folding mechanism. Owing to the absence of available anti-EspP passenger antibodies for pulse-chase analysis, culture supernatant fractions from TOP10 and BW25113 $\Delta$ ompT cells expressing EspP and EspP $\Delta$ L5 were analyzed for the presence of cleaved passenger domain. SDS-PAGE demonstrated that the EspP passenger secreted from BW25113 $\Delta$ ompT accumulated in the culture supernatant as one distinct fragment, whereas a breakdown product was observed when secreted from TOP10 (Fig. 3b, top panel). This breakdown product has been observed previously when EspP was over-expressed from a high-copy plasmid in an *ompT* + *E. coli* strain<sup>42</sup>. In contrast, SDS-PAGE demonstrated that the EspP $\Delta$ L5 passenger secreted from both strains accumulated in the culture supernatant at negligible levels, despite the presence of cleaved EspP $\Delta$ L5  $\beta$ -barrels in whole cell lysates (Fig. 3b, bottom panel). Together, these data suggest that the L5  $\beta$ -hairpin also mediates folding of the EspP passenger domain.

Although the conservation of residues within L5 of Pet and EspP is consistent with a conserved mechanism of passenger domain folding, the question becomes why is the sequence so conserved if it can be varied? To address this question, we performed a multiple sequence alignment using Muscle<sup>43</sup> on the amino acid sequence of SPATE and non-SPATE  $\beta$ -barrel domains that are representative of 11 distinct Types of  $\beta$ -barrel domains<sup>6</sup> (Supplementary Fig. 6a). These distinct  $\beta$ -barrel Types are categorized based on the arrangement of conserved motifs and phylogenetic clustering<sup>6</sup>. Identity between SPATE  $\beta$ -barrel domains, which are all Type 5  $\beta$ -barrel domains, is high, ranging from 59.21% to 99.64% (Supplementary Fig. 6b). From this, it follows that L5 is the same length in all SPATE  $\beta$ -barrel domains and contains several residues that are highly or completely conserved. In contrast, residues within L5 of non-SPATE  $\beta$ -barrel domains differ substantially with each other and with those in SPATE  $\beta$ -barrel domains (Supplementary Fig. 6a). Percent identity between these  $\beta$ -barrel domains ranges from 11.64% to 37.46% and even those belonging to the same  $\beta$ -barrel domain



type share low sequence identity (Supplementary Fig. 6c). Together, these data suggest that the conservation of residues within L5 in SPATE  $\beta$ -barrel domains could be reflective of the high conservation of the  $\beta$ -barrel domains among the SPATE subfamily and vice versa for non-SPATE  $\beta$ -barrel domains.

Next, we examined structures of solved non-SPATE  $\beta$ -barrel domains with electron density in the L5 region to determine whether L5 regions with low sequence identity to Pet or EspP still form a  $\beta$ -hairpin. This analysis revealed that the L5 region in the EstA<sup>7</sup> and AIDA-I<sup>44</sup>  $\beta$ -barrels both form a  $\beta$ -hairpin (Supplementary Fig. 7a, b). Although there is no structure for the AIDA-I passenger domain, it is predicted to be structurally homologous to that of Ag43, a  $\beta$ -helical passenger domain that mediates biofilm formation in *E. coli* through self-association and cell aggregation<sup>45–47</sup>. In contrast, the EstA passenger domain has a globular fold where a central four-stranded parallel  $\beta$ -sheet sits in close proximity to the L5  $\beta$ -hairpin<sup>7</sup> (Supplementary Fig. 7a). To test whether the EstA and AIDA-I  $\beta$ -hairpins have the potential to promote folding, we replaced the residues in the Pet L5  $\beta$ -hairpin with residues D<sub>556</sub>–L<sub>561</sub> and L<sub>564</sub>–E<sub>572</sub> from loop 5 from EstA, and with residues G<sub>1227</sub>–L<sub>1237</sub> from loop 5 of AIDA-I to create Pet<sup>L5EstA</sup> and Pet<sup>L5AIDA-I</sup>, respectively (Supplementary Fig. 7c). Similar to L5 in Pet, the first  $\beta$ -strand in the EstA  $\beta$ -hairpin is six residues in length, whereas that of the AIDA-I  $\beta$ -hairpin is shorter and only four residues in length. Pulse-chase expression assays showed that there was no obvious difference in the folding of Pet<sup>L5EstA</sup> relative to that observed in wild type Pet (Supplementary Fig. 7d). When grafted onto the Pet  $\beta$ -barrel domain, the short AIDA-I  $\beta$ -hairpin was also able to support folding of the Pet passenger domain. Although it is unclear why the first  $\beta$ -strand in the AIDA-I  $\beta$ -hairpin, but not the OmpF  $\beta$ -hairpin, is able to promote folding of the Pet passenger domain, despite both  $\beta$ -strands being four residues in length, it supports the notion that  $\beta$ -strand propensity and length might play a role in passenger folding. Nevertheless, we can conclude that while sequence alignments provide little insight into whether L5-assisted passenger folding might be conserved, studies with grafted heterologous loops coupled with structure-based analysis make way for a conserved role for L5 in the folding of SPATE and non-SPATE passenger domains, regardless of whether the passenger fold is  $\beta$ -helical or globular with mixed  $\alpha$ / $\beta$ -content.

### Passenger folding occurs without cellular machinery in vitro.

To investigate whether L5 of Pet is critical for folding of the passenger domain independently of the cellular machinery, the folding of Pet was monitored directly and in isolation. To this end, we engineered plasmids for the overexpression of proteins Pet <sup>$\Delta$ 1–554</sup> and Pet $\Delta$ L5 <sup>$\Delta$ 1–554</sup>, comprising the wild-type and L5-truncated  $\beta$ -barrel, respectively, and the 464 residues (~51 kDa) of the passenger located immediately N-terminal to the  $\beta$ -barrel domain (Fig. 4a). Inclusion bodies containing Pet <sup>$\Delta$ 1–554</sup> and Pet $\Delta$ L5 <sup>$\Delta$ 1–554</sup> were solubilized in urea and refolded in vitro by rapid dilution of the denaturant into detergent micelles as described previously<sup>34</sup>. The efficiency of the refolding reaction was monitored by the appearance of ~30 kDa and ~51 kDa cleavage products corresponding to the  $\beta$ -barrel and released passenger domains, respectively, from autocatalytic processing of the full-length ~81 kDa proteins (Fig. 4b).

To compare the folding of the Pet <sup>$\Delta$ 1–554</sup> and Pet $\Delta$ L5 <sup>$\Delta$ 1–554</sup> passenger domains, a time course trypsin assay was used to generate unique tryptic fingerprints caused by the transient exposure of trypsin-accessible arginine and lysine residues during passenger folding. In this assay, aliquots of the reaction mixture, at increasing incubation times, were treated with trypsin and then

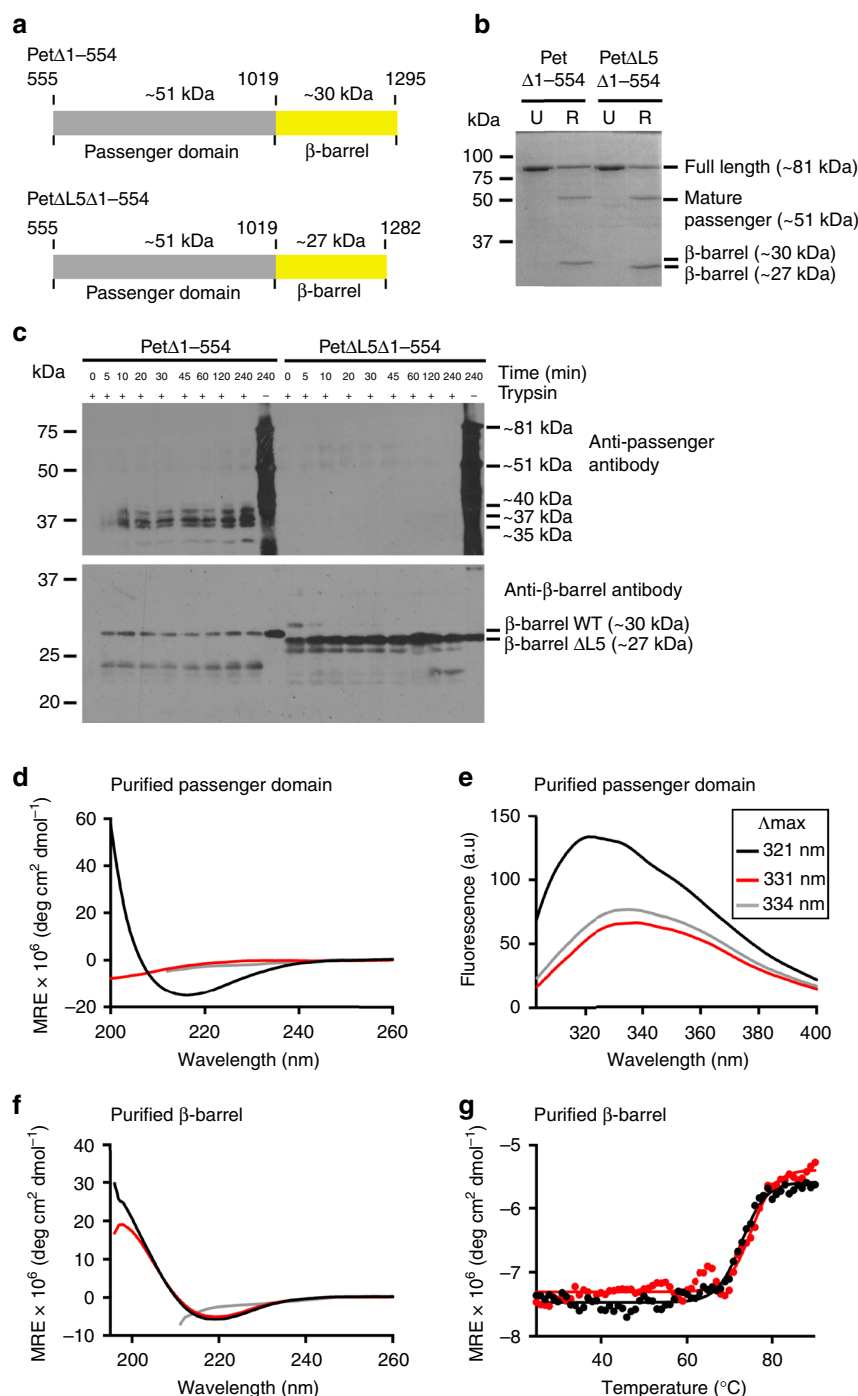
mixed with SDS to stop folding prior to SDS-PAGE analysis and immunoblotting using anti-passenger antibody (Fig. 4c, top panel). The folded Pet <sup>$\Delta$ 1–554</sup> passenger was largely resistant to trypsinolysis, generating trypsin-resistant fragments between ~35 and ~40 kDa, whereas the Pet $\Delta$ L5 <sup>$\Delta$ 1–554</sup> passenger was degraded completely. The presence of cleaved  $\beta$ -barrels in duplicate trypsin treated samples, detected using anti- $\beta$ -barrel antibody (Fig. 4c, bottom panel), indicated that truncation of L5 does not inhibit autocatalytic release of the passenger into the refolding buffer. Thus, we purified the Pet <sup>$\Delta$ 1–554</sup> and Pet $\Delta$ L5 <sup>$\Delta$ 1–554</sup> passengers isolated from the refolding buffer (Supplementary Fig. 8a) and characterized their fold using spectroscopic methods. The far-UV circular dichroism (CD) spectrum demonstrated that only the passenger derived from Pet <sup>$\Delta$ 1–554</sup> acquired secondary structure after refolding, whereas the spectrum of the Pet $\Delta$ L5 <sup>$\Delta$ 1–554</sup> passenger appeared almost indistinguishable from that of the unfolded protein in guanidine hydrochloride (Fig. 4d). Similarly, the tryptophan fluorescence spectrum of the passenger derived from Pet <sup>$\Delta$ 1–554</sup> demonstrated a maximum emission wavelength ( $\lambda_{\text{max}}$ ) of 321 nm, whereas the spectrum of the Pet $\Delta$ L5 <sup>$\Delta$ 1–554</sup> passenger appeared almost indistinguishable from that of the unfolded protein in guanidine hydrochloride, with a red-shifted  $\lambda_{\text{max}}$  of 331 nm (Fig. 4e). Together, these data suggest that the Pet  $\beta$ -barrel is a folding vector where L5 assists folding possibly by providing a structural template for the nucleation of passenger folding, a process that occurs independently of the cellular machinery in vitro.

**Passenger misfolding is not caused by a misfolded  $\beta$ -barrel.** As autocatalytic processing of SPATE passengers requires a folded  $\beta$ -barrel domain<sup>29,31,32</sup>, the presence of cleaved Pet <sup>$\Delta$ 1–554</sup> and Pet $\Delta$ L5 <sup>$\Delta$ 1–554</sup>  $\beta$ -barrels in trypsin-treated samples (Fig. 4c, bottom panel) also suggested that truncation of L5 does not inhibit folding of this domain in vitro. To validate these findings, we analyzed the structure of the refolded, cleaved and purified Pet <sup>$\Delta$ 1–554</sup> and Pet $\Delta$ L5 <sup>$\Delta$ 1–554</sup>  $\beta$ -barrel domains (Supplementary Fig. 8b) by CD spectroscopy. CD spectra of both the Pet <sup>$\Delta$ 1–554</sup> and Pet $\Delta$ L5 <sup>$\Delta$ 1–554</sup>  $\beta$ -barrels showed a characteristic  $\beta$ -signature with a minimum at 218 nm (Fig. 4f). Furthermore, CD collected as a function of temperature where denaturation of  $\beta$ -sheet structure is monitored in 1% SDS at 218 nm, showed that the apparent thermal stability of the Pet $\Delta$ L5 <sup>$\Delta$ 1–554</sup>  $\beta$ -barrel is indistinguishable from that of Pet <sup>$\Delta$ 1–554</sup> (Fig. 4g). Together, these data confirm that both  $\beta$ -barrel domains are folded correctly and strongly suggest that the folding defect observed for the passenger domain derived from Pet $\Delta$ L5 <sup>$\Delta$ 1–554</sup> is a direct consequence of the L5-truncation.

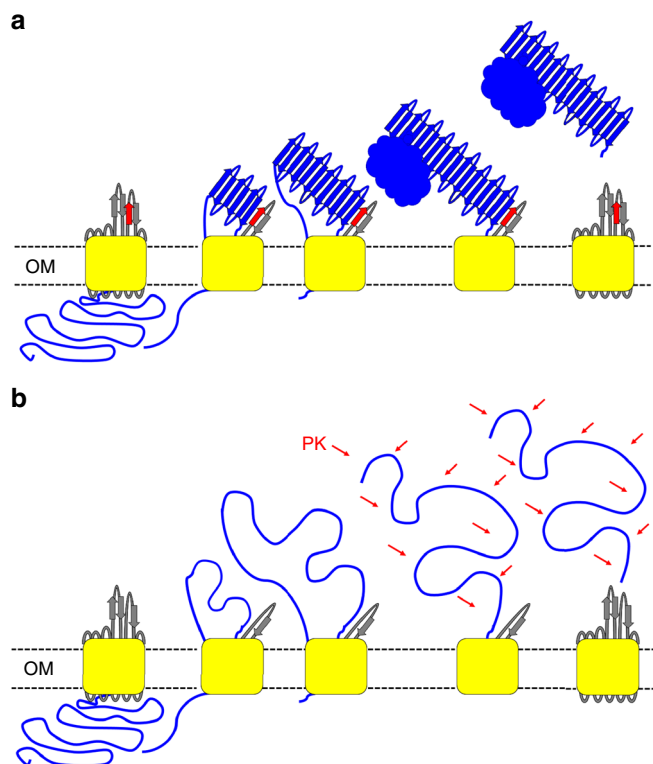
### Discussion

How autotransporter passenger domains fold into a  $\beta$ -helix in the absence of a chemical energy source or electrochemical gradient has remained controversial. In this study, we analyzed the folding of Pet in bacterial cells and in vitro. We find that the first  $\beta$ -strand in L5 is required to efficiently and rapidly fold the passenger domain into its native, protease-resistant structure, suggesting that it acts as a scaffold for passenger folding in a manner that is reminiscent of  $\beta$ -strand augmentation.  $\beta$ -strand augmentation is a process where non-covalent protein–protein interactions in a variety of biological processes are mediated by the addition of a  $\beta$ -strand in an unstructured ligand to a  $\beta$ -strand or  $\beta$ -sheet in a preformed, independently folded receptor, resulting in the formation of a continuous  $\beta$ -sheet<sup>48</sup>. For example,  $\beta$ -strand augmentation is one of the proposed mechanisms by which BamA mediates folding of OMP substrates<sup>49–51</sup>. Furthermore, the observation that the length and  $\beta$ -strand propensity of the very





**Fig. 4** Truncation of L5 affects folding of the Pet passenger domain in vitro. **a** Schematic of the Pet $\Delta 1-554$  and Pet $\Delta L5\Delta 1-554$  proteins. The 464-residue passenger incorporates residues 555 to 1018 of Pet and includes the  $\alpha$ -helical segment, which connects this domain to the wild-type (277 residues; 1019 to 1295) or L5-truncated (264 residues; 1019 to 1282)  $\beta$ -barrel domain. Cleavage within the  $\alpha$ -helical segment occurs between residues N<sub>1018</sub> and N<sub>1019</sub>, with this autocatalytic cleavage converting the ~81 kDa unfolded, full-length protein into ~51 kDa and ~30/27 kDa proteins corresponding to the cleaved passenger and  $\beta$ -barrel domains, respectively. **b** Cells expressing Pet $\Delta 1-554$  and Pet $\Delta L5\Delta 1-554$  were collected after induction with IPTG. These proteins were recovered from inclusion bodies, solubilized in urea, and then refolded by rapid dilution into LDAO detergent micelles. The refolding reaction was monitored by SDS-PAGE and Coomassie blue staining of the gels. Processing of the truncated passenger domain resulted in the autocatalytic conversion of the ~81 kDa full-length protein to ~51 kDa and ~30/27 kDa proteins, and served as an indicator of  $\beta$ -barrel folding. U, unfolded protein; R, refolded protein. **c** Folding of Pet $\Delta 1-554$  (left) and Pet $\Delta L5\Delta 1-554$  (right) monitored by the accumulation of trypsin-resistant fragments. In this assay, folding was initiated and aliquots of the reaction mixture, at increasing incubation times, were treated with trypsin and then mixed with SDS to stop folding. Samples were analyzed by SDS-PAGE and immunoblotting using anti-passenger domain (top panel) and anti- $\beta$ -barrel domain (bottom panel) antibodies. Far-UV CD (**d**) and tryptophan fluorescence (**e**) spectra of unfolded protein (grey), and in vitro folded, cleaved, and purified Pet $\Delta 1-554$  (black) and Pet $\Delta L5\Delta 1-554$  (red) passenger domains. **f** Far-UV CD spectra of unfolded protein (grey), and in vitro folded, cleaved and purified Pet $\Delta 1-554$  (black) and Pet $\Delta L5\Delta 1-554$  (red)  $\beta$ -barrel domains. **g** Thermal denaturation of in vitro folded, cleaved and purified Pet $\Delta 1-554$  (black) and Pet $\Delta L5\Delta 1-554$  (red)  $\beta$ -barrel domains monitored in 1% SDS at 218 nm. MRE; mean residue ellipticity



**Fig. 5** Proposed mechanism for passenger domain folding. **a** Schematic for folding shows that the first  $\beta$ -strand in L5 (red) interacts with the stable core in an antiparallel fashion during translocation, thereby polymerizing processive in-register stacking of the  $\beta$ -rungs and formation of the  $\beta$ -helix. Once the extreme N-terminus of the passenger domain has traversed the  $\beta$ -barrel pore, an autocatalytic cleavage reaction within an intra-barrel  $\alpha$ -helical segment liberates the ~106 kDa mature form of the passenger domain, leaving the ~30 kDa  $\beta$ -barrel domain embedded in the outer membrane. **b** Schematic shows that in the absence of the L5  $\beta$ -hairpin, the potential for  $\beta$ -strand augmentation is eliminated and thus interactions between the passenger domain and the first  $\beta$ -strand in L5 are prevented, thereby perturbing the formation of a continuous  $\beta$ -helix. Nevertheless, once the extreme N-terminus of the passenger domain has traversed the  $\beta$ -barrel pore, autocatalytic cleavage liberates the ~106 kDa mature form of the passenger domain, leaving the ~30 kDa  $\beta$ -barrel domain embedded in the outer membrane. As the passenger domain adopts an unfolded conformation, it is susceptible to degradation by host and exogenous proteases

C-terminal end of SurA, but not its specific amino acid sequence, are important for SurA function<sup>52</sup>, led to the proposition that the chaperone function of SurA also necessitates  $\beta$ -strand augmentation with its substrates<sup>50</sup>.

Although we have not yet identified the region of the passenger that transiently interacts with L5, we hypothesize that this region lies within the stable core given that these C-terminal  $\beta$ -helical rungs have been shown to appear first at the cell surface<sup>3,19,23,24</sup>. In support of this hypothesis are studies showing that mutation or deletion of the C-terminal  $\beta$ -helical rungs inhibits their folding<sup>19</sup> or the folding of the passenger N-terminus<sup>37</sup>, respectively. We propose that the first  $\beta$ -strand in L5 hydrogen bonds with the C-terminal portion of the unfolded passenger domain in an anti-parallel manner, which initiates polymerization of the stable core into a stack of in-register  $\beta$ -helical rungs to which the remaining N-terminal  $\beta$ -helical backbone is built upon (Fig. 5a). Our observation that the  $\beta$ -strand propensity of the L5  $\beta$ -hairpin, but not its specific amino acid sequence is important

for mediating folding of the Pet passenger domain suggests that the function of the L5  $\beta$ -hairpin requires  $\beta$ -augmentation with the passenger domain, as opposed to recognition of a specific sequence. This conclusion is consistent with the greater amino acid sequence diversity found in SPATE passenger domains, which have amino acid identities ranging from 35% to 55%<sup>42</sup>. In the absence of the L5  $\beta$ -hairpin, we find that the passenger domain mainly adopts an unfolded conformation and is therefore susceptible to degradation by host and exogenous proteases. We propose that when the potential for  $\beta$ -strand augmentation is eliminated, hydrogen bonding between the unfolded passenger and the first  $\beta$ -strand in L5 is prevented, and that it is the absence of these interactions at the very beginning of the nucleation process that perturbs the formation of a continuous  $\beta$ -helix (Fig. 5b).

Crystal structures of EspP show that L5 can adopt two different conformations<sup>29,30</sup>. The crystal structure of the EspP  $\beta$ -barrel domain post cleavage of the passenger domain (PDB code 2QOM) shows L5 in a 'closed' conformation, folded into the pore of the  $\beta$ -barrel where it hydrogen bonds with several  $\beta$ -strands. In contrast, the crystal structure of the EspP  $\beta$ -barrel domain with the most C-terminal portion of the passenger domain still attached (PDB code 3SLJ) shows L5 in an 'open' conformation, projecting into the extracellular space. Here, the uncleaved passenger domain traverses the length of the pore, thereby physically blocking L5 from making extensive contact with the  $\beta$ -barrel. Instead, the second  $\beta$ -strand in L4 interacts with the first  $\beta$ -strand in L5, forming a four-stranded  $\beta$ -sheet (Fig. 1b). Given the conformational flexibility of L5, we speculate that the L4 and L5 interaction is transient and that the  $\beta$ -hairpins separate during passenger domain translocation to allow L5 to interact with the passenger domain. In this scenario, L4 functions to restrain L5 movement and ensures that L5 is correctly positioned to interact with the passenger during translocation, thereby indirectly promoting folding of this domain. However, translocation occurs non-optimally through the L4-truncated  $\beta$ -barrel domain, suggesting a role for L4 in streamlining efficient passenger translocation. Hence, in an alternative scenario, L4 feeds the passenger domain to L5. Certainly, L4 forms part of a hydrophobic cavity in the BrkA  $\beta$ -barrel domain (which is also present in EspP) that is required to efficiently incorporate the passenger domain in a hairpin conformation, thereby promoting fast and efficient translocation through the  $\beta$ -barrel lumen<sup>35</sup>.

L5-assisted folding of the passenger domain can explain the molecular basis for folding defective phenotypes observed in previous studies that investigated the in vitro folding of passengers in isolation from their cognate  $\beta$ -barrels. Studies on pertactin, a classical autotransporter from *B. pertussis*, demonstrated that although folding of the passenger N terminus requires the presence of the passenger C terminus<sup>53</sup>, a folded C-terminus is not sufficient to promote processive folding of the N terminus<sup>3,54</sup>. It was further shown that refolding of the pertactin passenger into a  $\beta$ -helix in isolation is a slow (> 10 h), concerted and multistep process involving the entire protein, rather than vectorial<sup>3,54</sup>. By contrast, autotransporter assembly in vivo occurs rapidly (~10 min)<sup>18</sup> and is in line with the typical doubling time of *E. coli* (~20 min). We conclude that the mechanistic differences between passenger folding with and without the  $\beta$ -barrel have transpired due to the inability to recapitulate in vitro the effects of the autotransporter  $\beta$ -barrel on passenger domain folding in the bacterial cell.

In addition to actively augmenting folding, as proposed previously<sup>25</sup>, it is also feasible that the  $\beta$ -barrel accelerates passenger folding through promotion of vectorial folding during translocation. This would allow the passenger C terminus to explore its conformational space before the N terminus has been

translocated, thereby decreasing the probability of populating non-native structures. Such non-native structures likely cause slow in vitro refolding kinetics because they require extensive reorganization before the native conformation can be reached<sup>54</sup>. Vectorial folding would also ensure that the correct part of the passenger is able to interact with the first  $\beta$ -strand in L5 at the very beginning of the translocation process.

In summary, we have discovered that the SPATE  $\beta$ -barrel domain is a folding vector that can act independently of the cellular machinery to promote folding of the passenger domain on a biologically relevant timescale and we provide evidence to show that the efficiency of this process is dependent on the L5  $\beta$ -hairpin. The significance of our work is evident in the inability of the passenger domain to fold into its biologically active form in the absence of its interaction with L5, a phenotype that could be exploited for therapeutic purposes.

## Methods

**Reagents and bacterial strains.** Bacteria were grown at 37 °C in Luria–Bertani broth and where necessary, the growth medium was supplemented with 30  $\mu\text{g mL}^{-1}$  kanamycin and/or 100  $\mu\text{g mL}^{-1}$  ampicillin, 0.2% (vol/vol) D-glucose, 0.02% (vol/vol) L-arabinose, or 0.5 mM isopropylthio- $\beta$ -galactoside. Details of the *E. coli* strains (TOP10, BW25113, and BL21 [DE3]) and plasmids used in this study are presented in Supplementary Table 1, and primers used for plasmid construction are presented in Supplementary Table 2.

**Plasmid construction.** pBADPet has been described previously<sup>34</sup>. pBADEspP, pBADEspPAL5, and all pBADPet derivatives except PetL5 $\beta$ 1/P were synthesized de novo by GenScript. To construct PetL5 $\beta$ 1/P, megaprimer PCR was performed as described previously<sup>34,55</sup>. Briefly, the round 1 PCR was performed using 500 ng of template DNA (pBADPet) with 1  $\mu\text{g mL}^{-1}$  each of primers MPL5 $\beta$ 1Pro and HindIII/PetRv. The round 2 PCR was performed with 4  $\mu\text{g}$  of megaprimer and 1  $\mu\text{g}$  of primer SalIPetFw using 500 ng of template DNA (pBADPet). The resulting amplicon and target vector (pBADPet) were then digested with HindIII and SalI and ligated. To construct pETPet $\Delta$ L5<sup>554</sup> and pETPet $\Delta$ L5<sup>554</sup> the Pet  $\beta$ -barrel domain and the last 464 residues of the Pet passenger domain were amplified using 500 ng of template DNA (pBADPet and pBADPet $\Delta$ L5) with 1  $\mu\text{g mL}^{-1}$  each of primers NdeIPet464Fw and XhoIPetRv. The subsequent amplicons and target vector (pET-22b+) were then digested with NdeI and XhoI, and ligated for an in frame C-terminal hexahistidine-tag fusion. All DNA modifications were confirmed by sequencing.

**In vivo protein expression assays.** The pBAD expression system<sup>56</sup> was used to express protein as previously described<sup>34</sup>. In all cases, bacterial cultures were grown to an OD<sub>600</sub> of ~0.6 in media containing 0.2% glucose to repress protein synthesis prior to harvesting (2500  $\times$  g, 5 min, 4 °C). For pulse-chase expression of the pBADPet derivatives, the bacterial cells were resuspended in media containing 0.02% arabinose for five minutes (pulse) to induce Pet expression, collected, and then resuspended in media containing 0.2% glucose to “chase” the protein produced during the “pulse” phase of the experiment through the biogenesis pathway. Throughout the pulse-chase assay, three 1 mL aliquots were removed from each culture at each time point. The first aliquot was added to a tube containing trichloroacetic acid (TCA; final concentration, 10% [wt/vol]), and placed on ice. The second aliquot was added to a tube containing proteinase K (final concentration, 200  $\mu\text{g mL}^{-1}$ ), and placed on ice for 20 min to digest Pet exposed on the cell surface. The protease reaction was stopped by the addition of 2 mM phenylmethanesulfonyl fluoride (PMSF) before TCA precipitation. All TCA precipitated samples were washed with acetone, dried, and resuspended in 100  $\mu\text{L}$  of SDS-PAGE loading buffer per 1 OD<sub>600</sub> units of cells (third aliquot). Where necessary, a fourth aliquot was treated with 200  $\mu\text{g mL}^{-1}$  polymyxin B for 10 min on ice prior to proteinase K treatment and TCA precipitation. Ten microliters of the normalized samples were separated by SDS-PAGE on 3–14% gradient gels and subjected to immunoblotting using anti-Pet passenger domain antibodies at a 1:5,000 dilution. These antibodies have been described and validated previously<sup>28,34</sup>.

Bacterial cultures harbouring pBADEspP and pBADEspPAL5 were resuspended in arabinose for 30 min to induce EspP and EspPAL5 expression. The OD<sub>600</sub> of cultures were normalized (to permit comparison of secreted protein levels) and pelleted before filtering of the supernatant fractions through 0.22  $\mu\text{m}$  pore-size filters. A 10% (w/v) final concentration of TCA was then used to precipitate the secreted proteins as described above. Samples were separated by 10% SDS-PAGE and detected by staining with Coomassie Brilliant Blue R250. Whole cell lysates were separated by SDS-PAGE on 3–14% gradient gels and subjected to immunoblotting using anti-Pet  $\beta$ -barrel domain antibodies at a 1:5,000 dilution as described above. The Pet and EspP  $\beta$ -barrel domains are so similar that these

antibodies cross-react with the EspP  $\beta$ -barrel domain. These antibodies have been described and validated previously<sup>34,57</sup>.

**Heat modifiability assay.** After pulse-chase protein expression, cells were harvested and prepared for heat modifiability experiments as described previously<sup>34</sup>. Briefly, 40  $\mu\text{g}$  of total membranes were mixed with SDS-PAGE loading buffer and then heated at temperatures between 25 °C and 90 °C. Samples were separated by SDS-PAGE on 3–14% gradient gels and subjected to immunoblotting using anti-Pet  $\beta$ -barrel domain antibodies as described above.

**Protein refolding and purification.** Protein expression and refolding were carried out as previously described<sup>34</sup>. Briefly, *E. coli* BL21 (DE3) cells transformed with pETPet $\Delta$ L5<sup>554</sup> and pETPet $\Delta$ L5<sup>554</sup> were harvested (4200  $\times$  g, 5 min, 4 °C) after a 4 h (pETPet $\Delta$ L5<sup>554</sup>) or 18 h (pETPet $\Delta$ L5<sup>554</sup>) induction with 0.5 mM IPTG. Inclusion bodies were isolated by centrifugation (10,000  $\times$  g, 15 min, 4 °C) and solubilized in 8 M urea. Pet $\Delta$ L5<sup>554</sup> and Pet $\Delta$ L5<sup>554</sup> were then refolded from the urea extracts by rapid 10-fold dilution into tris-buffered saline containing 0.5 % (w/v) lauryldimethylamine-oxide (LDAO) detergent micelles at 35 °C. Affinity purification using Ni-NTA resin was used to separate the refolded passenger and  $\beta$ -barrel domains. Addition of Bio-Beads<sup>TM</sup> SM-2 Resin (BioRad) at 0.1 g mL<sup>-1</sup> to the unbound flow-through containing the cleaved passenger domain removed the LDAO detergent and residual  $\beta$ -barrel. Subsequent gel filtration chromatography in 50 mM Tris pH 8.0, 150 mM NaCl (Superdex s200-10/300, GE HealthCare) resulted in purification of the ~51 kDa passenger domain. Bound  $\beta$ -barrel was eluted from the Ni-NTA resin and further purified by gel filtration chromatography in 50 mM Tris pH 8.0, 150 mM NaCl, 0.05% LDAO (wt/vol) (Superdex s200-10/300, GE HealthCare), which resulted in purification of the folded ~30/27 kDa forms.

**Biophysical analysis of protein structure.** Kinetics of trypsin cleavage profiles during Pet $\Delta$ L5<sup>554</sup> and Pet $\Delta$ L5<sup>554</sup> refolding were initiated as described above and were followed by limited treatment with trypsin, which was performed as described previously<sup>34</sup>. Briefly, aliquots of the reaction mixture were removed at the time points indicated and treated with trypsin (20  $\mu\text{g mL}^{-1}$ ) for 20 min on ice before quenching digestion and folding with 2 mM PMSF and SDS loading buffer, respectively. The resulting trypsin fragments were separated by SDS-PAGE and subjected to immunoblotting using anti-passenger and anti- $\beta$ -barrel domain antibodies as described above.

Far-UV CD spectra were collected from 190 to 260 nm using a Chirascan spectrometer. Samples containing 0.1 mg mL<sup>-1</sup> of purified protein were measured at room temperature with a 1 mm cuvette, 1.0 nm bandwidth, 1 s integration time, and 20 nm/min scanning speed. Three scans were averaged, and the spectra were corrected for buffer contribution. For thermal denaturation measurements, the denaturation of  $\beta$ -sheet structure was monitored in the presence of 1% SDS at 218 nm as a function of temperature from 25 °C to 95 °C at a temperature change of 1 °C min<sup>-1</sup>. Thermal unfolding was irreversible. In both cases, the CD spectra were normalized to the mean residue molar ellipticity.

A Varian Cary Eclipse fluorimeter was used to collect emission spectra between 300–400 nm upon excitation of tryptophan at 295 nm. Samples containing 0.01 mg mL<sup>-1</sup> of purified protein were collected at room temperature with a 1 cm cuvette, 5 nm slit width, 0.1 s integration time, and an increment of 1.0 nm. Three scans were averaged, and the spectra were subtracted for buffer contribution.

**Data availability.** All data generated or analysed during this study are included in this published article and its Supplementary Information files. The raw image files of the cropped immunoblots and Coomassie-stained gels displayed in the Figures and Supplementary Figures can be found in Supplementary Fig. 9. Other data are available from the corresponding author upon reasonable request.

Received: 31 July 2017 Accepted: 27 February 2018

Published online: 11 April 2018

## References

1. Fan, E., Chauhan, N., Udatha, D. B. R. K. G., Leo, J. C. & Linke, D. Type V secretion systems in bacteria. *Microbiol. Spectr.* **4**, <https://doi.org/10.1128/microbiolspec.VMBF-0009-2015> (2016).
2. Drobnak, I. et al. Of linkers and autochaperones: an unambiguous nomenclature to identify common and uncommon themes for autotransporter secretion. *Mol. Microbiol.* **95**, 1–16 (2015).
3. Junker, M. et al. Pertactin  $\beta$ -helix folding mechanism suggests common themes for the secretion and folding of autotransporter proteins. *Proc. Natl Acad. Sci. USA* **103**, 4918–4923 (2006).



4. Kajava, A. V. & Steven, A. C. The turn of the screw: variations of the abundant  $\beta$ -solenoid motif in passenger domains of Type V secretory proteins. *J. Struct. Biol.* **155**, 306–315 (2006).
5. Leyton, D. L., Rossiter, A. E. & Henderson, I. R. From self sufficiency to dependence: mechanisms and factors important for autotransporter biogenesis. *Nat. Rev. Microbiol.* **10**, 213–225 (2012).
6. Celik, N. et al. A bioinformatic strategy for the detection, classification and analysis of bacterial autotransporters. *PLoS ONE* **7**, e43245 (2012).
7. van den Berg, B. Crystal structure of a full-length autotransporter. *J. Mol. Biol.* **396**, 627–633 (2010).
8. Henderson, I. R., Navarro-Garcia, F., Desvaux, M., Fernandez, R. C. & Ala'Aldeen, D. Type V protein secretion pathway: the autotransporter story. *Microbiol. Mol. Biol. Rev.* **68**, 692–744 (2004).
9. Dautin, N. Serine protease autotransporters of *Enterobacteriaceae* (SPATEs): biogenesis and function. *Toxins* **2**, 1179–1206 (2010).
10. Nishimura, K., Tajima, N., Yoon, Y.-H., Park, S.-Y. & Tame, J. H. Autotransporter passenger proteins: virulence factors with common structural themes. *J. Mol. Med.* **88**, 451–458 (2010).
11. Benz, I. & Schmidt, M. A. Structures and functions of autotransporter proteins in microbial pathogens. *Int. J. Med. Microbiol.* **301**, 461–468 (2011).
12. Selkig, J. et al. Discovery of an archetypal protein transport system in bacterial outer membranes. *Nat. Struct. Mol. Biol.* **19**, 506–510 (2012).
13. Shen, H.-H. et al. Reconstitution of a nanomachine driving the assembly of proteins into bacterial outer membranes. *Nat. Commun.* **5**, 5078 (2014).
14. Gruss, F. et al. The structural basis of autotransporter translocation by TamA. *Nat. Struct. Mol. Biol.* **20**, 1318–1320 (2013).
15. Ieva, R., Tian, P., Peterson, J. H. & Bernstein, H. D. Sequential and spatially restricted interactions of assembly factors with an autotransporter  $\beta$  domain. *Proc. Natl Acad. Sci. USA* **108**, E383–E391 (2011).
16. Ruiz-Perez, F. et al. Roles of periplasmic chaperone proteins in the biogenesis of serine protease autotransporters of *Enterobacteriaceae*. *J. Bacteriol.* **191**, 6571–6583 (2009).
17. Ruiz-Perez, F., Henderson, I. R. & Nataro, J. P. Interaction of FkpA, a peptidyl-prolyl cis/trans isomerase with EspP autotransporter protein. *Gut Microbes* **1**, 339–344 (2010).
18. Pavlova, O., Peterson, J. H., Ieva, R. & Bernstein, H. D. Mechanistic link between  $\beta$  barrel assembly and the initiation of autotransporter secretion. *Proc. Natl Acad. Sci. USA* **110**, E938–E947 (2013).
19. Peterson, J. H., Tian, P., Ieva, R., Dautin, N. & Bernstein, H. D. Secretion of a bacterial virulence factor is driven by the folding of a C-terminal segment. *Proc. Natl Acad. Sci. USA* **107**, 17739–17744 (2010).
20. Ieva, R. & Bernstein, H. D. Interaction of an autotransporter passenger domain with BamA during its translocation across the bacterial outer membrane. *Proc. Natl Acad. Sci. USA* **106**, 19120–19125 (2009).
21. Noinaj, N. et al. Structural insight into the biogenesis of  $\beta$ -barrel membrane proteins. *Nature* **501**, 385–390 (2013).
22. Gu, Y. et al. Structural basis of outer membrane protein insertion by the BAM complex. *Nature* **531**, 64–69 (2016).
23. Junker, M., Besingi, R. N. & Clark, P. L. Vectorial transport and folding of an autotransporter virulence protein during outer membrane secretion. *Mol. Microbiol.* **71**, 1323–1332 (2009).
24. Renn, J. P. & Clark, P. L. A conserved stable core structure in the passenger domain  $\beta$ -helix of autotransporter virulence proteins. *Biopolymers* **89**, 420–427 (2008).
25. Braselmann, E. & Clark, P. L. Autotransporters: the cellular environment reshapes a folding mechanism to promote protein transport. *J. Phys. Chem. Lett.* **3**, 1063–1071 (2012).
26. Baclayon, M. et al. Mechanical unfolding of an autotransporter passenger protein reveals the secretion starting point and processive transport intermediates. *ACS Nano* **10**, 5710–5719 (2016).
27. Sopova, Z. et al. A conserved aromatic residue in the autochaperone domain of the autotransporter Hbp is critical for initiation of outer membrane translocation. *J. Biol. Chem.* **285**, 38224–38233 (2010).
28. Eslava, C. et al. Pet, an autotransporter enterotoxin from enteroaggregative *Escherichia coli*. *Infect. Immun.* **66**, 3155–3163 (1998).
29. Barnard, T. J., Dautin, N., Lukacik, P., Bernstein, H. D. & Buchanan, S. K. Autotransporter structure reveals intra-barrel cleavage followed by conformational changes. *Nat. Struct. Mol. Biol.* **14**, 1214–1220 (2007).
30. Barnard, T. J. et al. Molecular basis for the activation of a catalytic asparagine residue in a self-cleaving bacterial autotransporter. *J. Mol. Biol.* **415**, 128–142 (2012).
31. Tajima, N., Kawai, F., Park, S.-Y. & Tame, J. R. H. A novel intein-like autoproteolytic mechanism in autotransporter proteins. *J. Mol. Biol.* **402**, 645–656 (2010).
32. Dautin, N., Barnard, T. J., Anderson, D. E. & Bernstein, H. D. Cleavage of a bacterial autotransporter by an evolutionarily convergent autocatalytic mechanism. *EMBO J.* **26**, 1942–1952 (2007).
33. Brunder, W., Schmidt, H. & Karch, H. EspP, a novel extracellular serine protease of enterohaemorrhagic *Escherichia coli* O157:H7 cleaves human coagulation factor V. *Mol. Microbiol.* **24**, 767–778 (1997).
34. Leyton, D. L. et al. A mortise–tenon joint in the transmembrane domain modulates autotransporter assembly into bacterial outer membranes. *Nat. Commun.* **5**, 4239 (2014).
35. Zhai, Y. et al. Autotransporter passenger domain secretion requires a hydrophobic cavity at the extracellular entrance of the  $\beta$ -domain pore. *Biochem. J.* **435**, 577–587 (2011).
36. Baba, T. et al. Construction of *Escherichia coli* K-12 in-frame, single-gene knockout mutants: the Keio collection. *Mol. Syst. Biol.* **2006**, 2006.0008 (2006).
37. Oliver, D. C., Huang, G., Nodel, E., Pleasance, S. & Fernandez, R. C. A conserved region within the *Bordetella pertussis* autotransporter BrkA is necessary for folding of its passenger domain. *Mol. Microbiol.* **47**, 1367–1383 (2003).
38. Ohnishi, S., Kamikubo, H., Onitsuka, M., Kataoka, M. & Shortle, D. Conformational preference of polyglycine in solution to elongated structure. *J. Am. Chem. Soc.* **128**, 16338–16344 (2006).
39. Bykov, S. & Asher, S. Raman studies of solution polyglycine conformations. *J. Phys. Chem. B* **114**, 6636–6641 (2010).
40. Khan, S., Mian, H. S., Sandercock, L. E., Chirgadze, N. Y. & Pai, E. F. Crystal structure of the passenger domain of the *Escherichia coli* autotransporter EspP. *J. Mol. Biol.* **413**, 985–1000 (2011).
41. Domingo Meza-Aguilar, J. et al. X-ray crystal structure of the passenger domain of plasmid encoded toxin (Pet), an autotransporter enterotoxin from enteroaggregative *Escherichia coli* (EAEC). *Biochem. Biophys. Res. Commun.* **445**, 439–444 (2014).
42. Dutta, P. R., Cappello, R., Navarro-García, F. & Nataro, J. P. Functional comparison of serine protease autotransporters of enterobacteriaceae. *Infect. Immun.* **70**, 7105–7113 (2002).
43. Edgar, R. C. MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res.* **32**, 1792–1797 (2004).
44. Gawarzewski, I. et al. Crystal structure of the transport unit of the autotransporter adhesin involved in diffuse adherence from *Escherichia coli*. *J. Struct. Biol.* **187**, 20–29 (2014).
45. Heras, B. et al. The antigen 43 structure reveals a molecular Velcro-like mechanism of autotransporter-mediated bacterial clumping. *Proc. Natl Acad. Sci. USA* **111**, 457–462 (2014).
46. Klemm, P., Hjerrild, L., Gjermansen, M. & Schembri, M. A. Structure-function analysis of the self-recognizing Antigen 43 autotransporter protein from *Escherichia coli*. *Mol. Microbiol.* **51**, 283–296 (2004).
47. Ulett, G. C. et al. Functional analysis of Antigen 43 in uropathogenic *Escherichia coli* reveals a role in long-term persistence in the urinary tract. *Infect. Immun.* **75**, 3233–3244 (2007).
48. Remaut, H. & Waksman, G. Protein–protein interaction through  $\beta$ -strand addition. *Trends Biochem. Sci.* **31**, 436–444 (2006).
49. Knowles, T. J. et al. Fold and function of polypeptide transport-associated domains responsible for delivering unfolded proteins to membranes. *Mol. Microbiol.* **68**, 1216–1227 (2008).
50. Zhang, X. C. & Han, L. How does a  $\beta$ -barrel integral membrane protein insert into the membrane? *Protein & Cell* **7**, 471–477 (2016).
51. Estrada Mallarino, L. et al. TtOmp85, a  $\beta$ -barrel assembly protein, functions by barrel augmentation. *Biochem. J.* **54**, 844–852 (2015).
52. Chai, Q. et al. Diverse sequences are functional at the C-terminus of the *E. coli* periplasmic chaperone SurA. *Protein Eng. Des. Sel.* **27**, 111–116 (2014).
53. Renn, J. P., Junker, M., Besingi, R. N., Braselmann, E. & Clark, P. L. ATP-independent control of autotransporter virulence protein transport via the folding properties of the secreted protein. *Chem. Biol.* **19**, 287–296 (2012).
54. Junker, M. & Clark, P. L. Slow formation of aggregation-resistant  $\beta$ -sheet folding intermediates. *Proteins* **78**, 812–824 (2010).
55. Leyton, D. L. et al. Size and conformation limits to secretion of disulfide-bonded loops in autotransporter proteins. *J. Biol. Chem.* **286**, 42283–42291 (2011).
56. Guzman, L. M., Belin, D., Carson, M. J. & Beckwith, J. Tight regulation, modulation, and high-level expression by vectors containing the arabinose PBAD promoter. *J. Bacteriol.* **177**, 4121–4130 (1995).
57. Rossiter, A. E. et al. The essential  $\beta$ -barrel assembly machinery complex components BamD and BamA are required for autotransporter biogenesis. *J. Bacteriol.* **193**, 4250–4253 (2011).

## Acknowledgements

We thank Alex Joule and Anthony Newman for technical assistance, Trevor Lithgow and Iain Hay for providing the *E. coli* BW25113 $\Delta$ ompT strain, John Carver for access to the Chirscan spectrometer, as well as Susan Howitt for critical comments on the manuscript. We gratefully acknowledge support of the Australian Research Council (ARC) for



research funding through the ARC Discovery Project grant DP160103294 (to D.L.L. and I.R.H.). M.A.S. is a National Health and Medical Research Council Senior Research Fellow (APP1106930), G.H.M.H. is supported by a Global Marie-Curie Fellowship (EU project 660083), and D.L.L. is an ARC Future Fellow (FT150100452).

### Author contributions

X.Y., M.D.J., G.H.M.H., I.R.H., and D.L.L. conceived and designed the experiments. X.Y., M.D.J., J.Z., and A.W.L. performed the experiments. X.Y., M.D.J., G.H.M.H., A.W.L., M.A.S., L.C.W., R.N.P., I.R.H., and D.L.L. analyzed the data. X.Y., M.D.J., and D.L.L. wrote manuscript. All authors have reviewed the manuscript.

### Additional information

**Supplementary Information** accompanies this paper at <https://doi.org/10.1038/s41467-018-03593-2>.

**Competing interests:** All authors declare no competing interests.

**Reprints and permission** information is available online at <http://npg.nature.com/reprintsandpermissions/>

**Publisher's note:** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2018